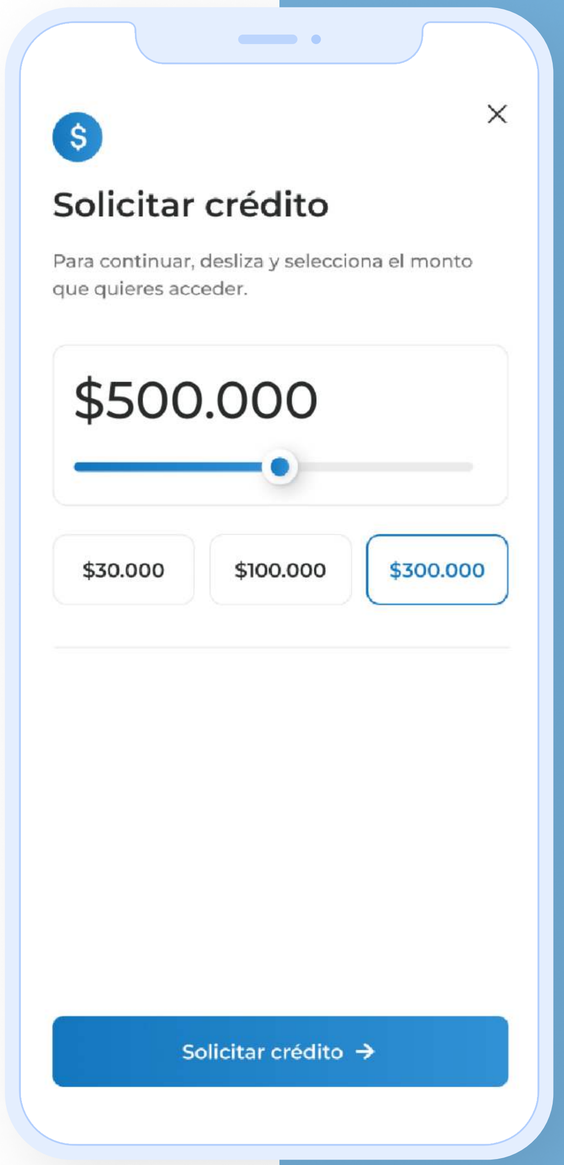

MONET
belvo.

belvo.

CASE STUDY **MONET**

How Monet built a credit score from open finance data

Monet and Belvo partnered to explore the use of open finance data **using machine learning techniques**, with the goal to explore the boundaries of scoring with alternative data. In this case study, we present the results of this collaboration – which show that standalone open finance data has sufficient predictive power for successful credit scoring.

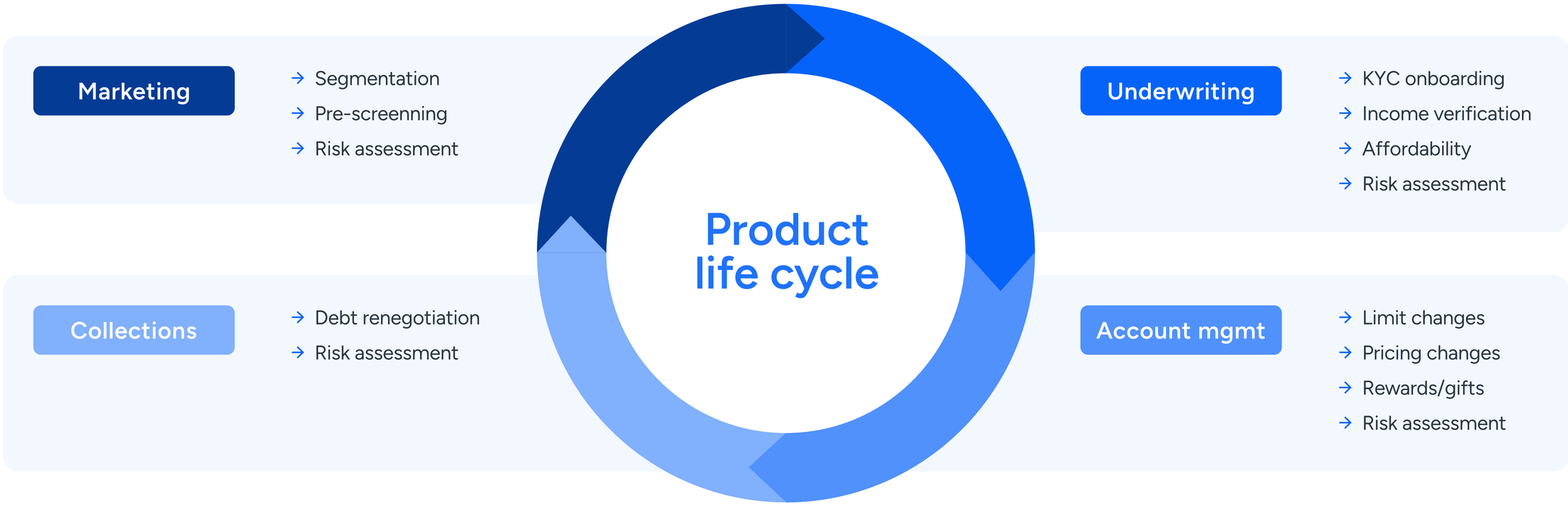


Context

Monet is a fintech specialized in small loans operating in Colombia with the mission to develop open finance-based inclusive financial services. Latin America's workers suffer from **low access to credit** and an expensive traditional financial system.

In the face of this, Monet's solution offers small loans to underserved customers. Monet has originated +570 K loans since June 2021. Their main pillar for loan approval since the beginning has been open finance data with rule-based systems thanks to their collaboration with Belvo. Monet does not use credit bureau information as their user base falls within the no-hit or thin file population.

Having gathered a significant amount of data over the years, Monet decided to **explore the use of probabilistic models** for new loan underwriting.



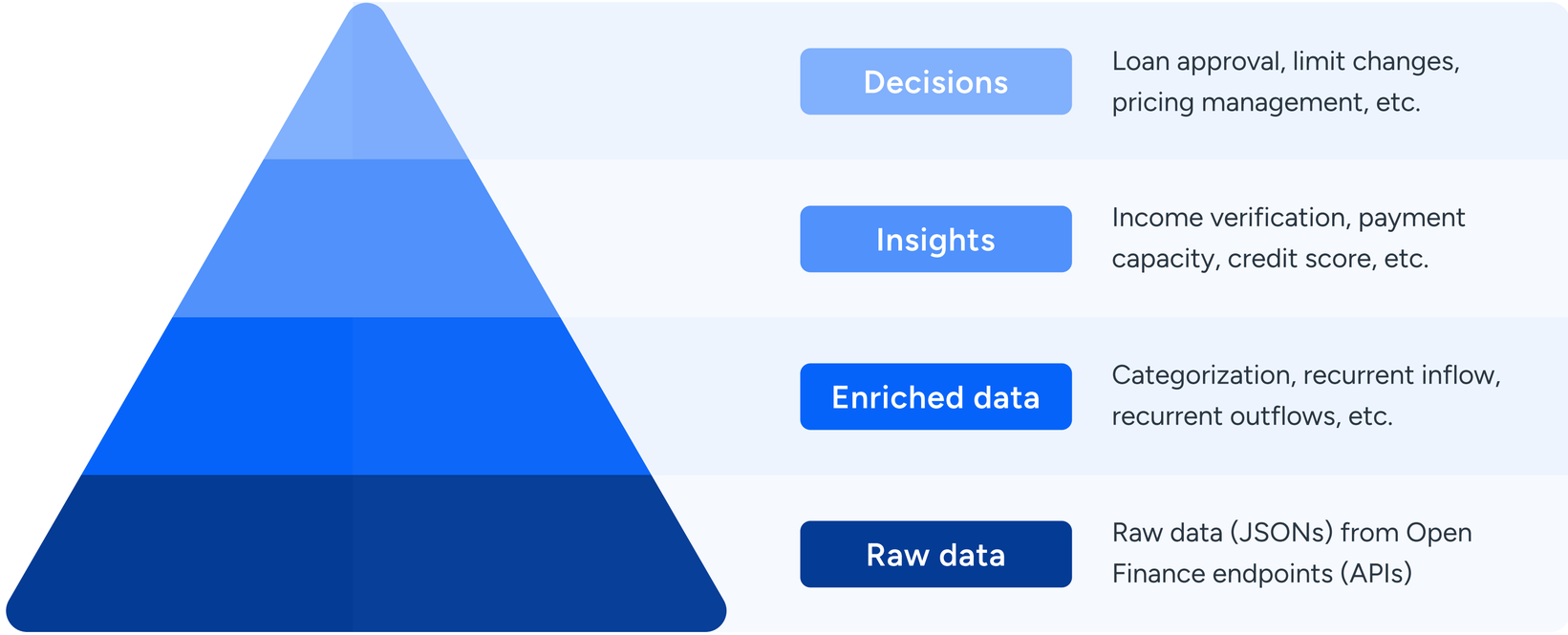
The project

The project started with data preparation. Belvo open finance data comes in a standardized schema as the API platform makes sure that the data is homogeneous across different sources. In order to expand feature space, a layer of enrichment on top of account and transaction information was applied.

This layer covered transaction categorization, recurrent income and recurrent expense verification.

Once the data was prepared, work continued with the creation of variables from enriched open finance data that could later be used for training and prediction (i.e. feature engineering). After various iterations, +350 features were created covering different time windows and types of open finance data. These features can be grouped into six classes:

- **Accounts.** Using information from saving accounts, loan accounts and credit cards.
- **Balances.** Built from accounts and transactions to reflect balance evolution.
- **Activity.** Considering activity patterns across value dates.
- **Transactionality.** Using the number and type of historic transactions.
- **Cashflow.** Using positive and negative cash-flows within given time windows.
- **Categories.** Combining transactions with categorization to build customer profiles.



Another important consideration was the **definition of good and bad payers** to be used in algorithm training. Given Monet's business model, an early payment default was considered as an appropriate performance time horizon.

For default definition, the level of days past due was chosen taking into consideration recovery curves from past cohorts.

With the final dataset in hand, the project moved into supervised learning.

The chosen machine learning classifier was XGBoost and the dataset was split into a train sample (80%) for hyperparameter and feature selection and a test sample (20%) for performance validation.

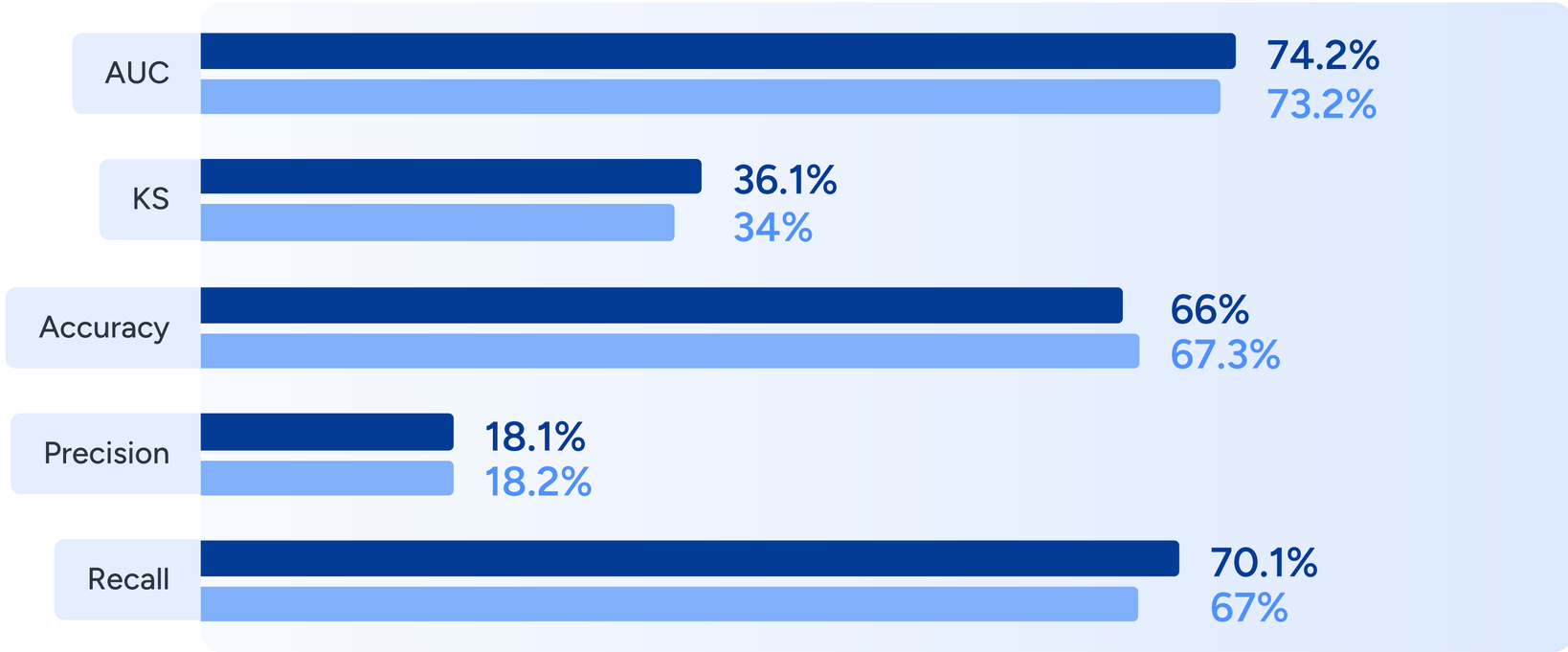
Hyperparameters were tuned using a randomized search configured for imbalanced classification and features were selected using SHAP values.

The results

From a classification point of view, the performance of the scoring model was considered successful with **AUC of 73% and KS of 34% on test sample**.

The AUC ('Area under the curve') and the KS (Kolmogorov-Smirnov test) are scoring quality measures to understand to what extent the model is able to distinguish between classes.

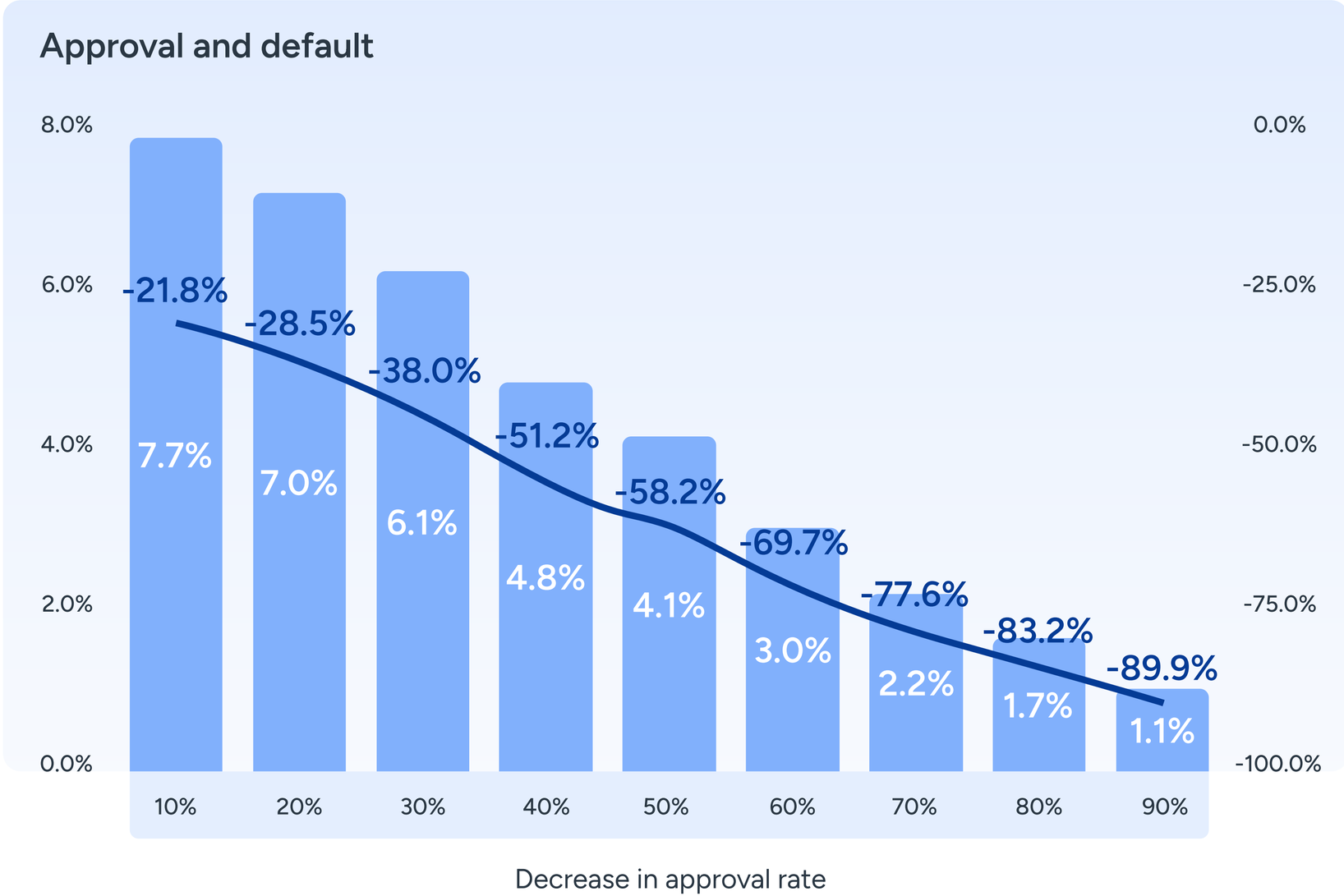
The graphic on the right shows common classification metrics for both train and test samples:



From a business point of view, the performance of the scoring model was also considered successful. **The company will save over 20%** in defaulted loans using probabilistic models.

The graph on the right shows the reduction in default rates that could be attained with changing thresholds on approval:

With these results in hand, Monet decided to **move forward and implement a probabilistic scoring model** for new loan underwriting. Open finance data will continue as the main pillar for decision making and as the amount of available data grows, going forward, Monet and Belvo will also explore the use of behavioral scores for other use cases.



belvo.

Are you looking to improve your credit scoring models?

Get in touch: hello@belvo.com